CSC321 Programming Assignment 3

Vitaly Topekha

March 23, 2018

1 Part 1. Encoder-Decoder Models and Capacity

1. Since the neural network has to compresss more of the necessary information into a fixed-length context vector, the architecrute won't perform well for longer sequences. That way, the decoder has less information to use for translation generation.

2. The model works great only for a specific kind of words, for instance with 1 or 2 characters for the first letter vowel words or fist letter consonant word. With 3 or more characters, translation is wrong. Words with consonant pairs like "tr" works well except for some specific cases of length like "truth". Consonant pairs of "sh" do not get translated well even for short length like "shy".

   Examples:
   i $->$ iway
   is $->$ isway
   she $->$ ethay
   a $->$ away
   no $->$ oway
   hello $->$ erlehay
   happy $->$ apalcay
   easy $->$ eayyway

2 Part 2. Teacher-Forcing

1. The model could be over-fit for the given training data. During the training time the model could receive different inputs from the test time, since the while training, the model uses previous output as an input. Therefore, the performance could be damaged. Sequence of previously generated samples diverges from sequence seen during training.

2. Diminish the gap between inpurs seen during training time and test time through randomly using generater output or ground truth from the previous step as input for the current time step.

3 Part 5. Attention Visualizations

The model does only work well for words with a common structure. If a words has a weird structre like "ssssssss" it doesn't work since there are not enough training examples for the cases, so it predicts it to something seen before

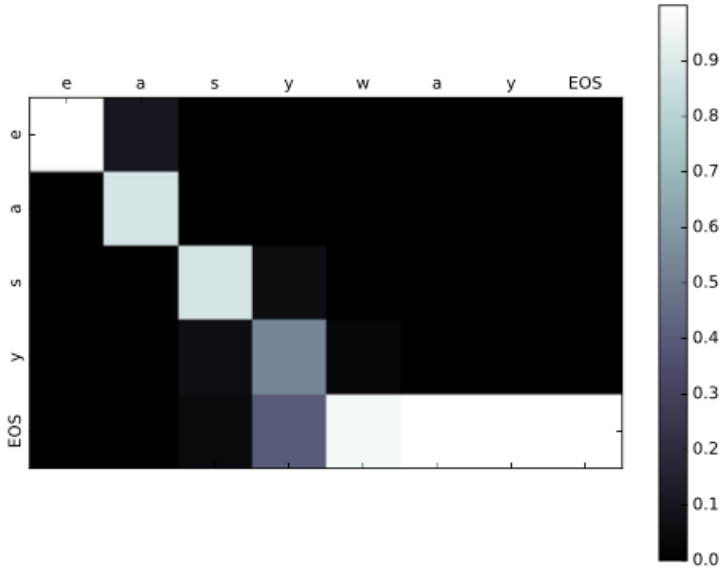

Figure 1: working: easy - easyway

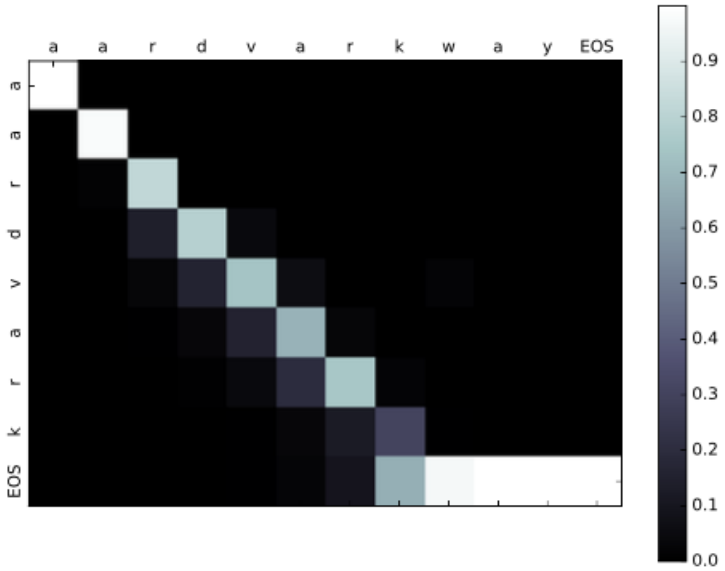

Figure 2: working: aardvark - aardvarkway

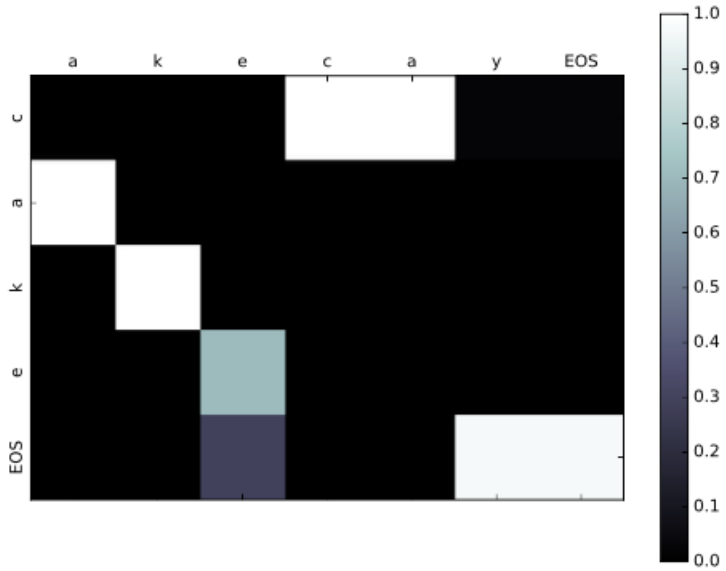

Figure 3: working: cake - akecay

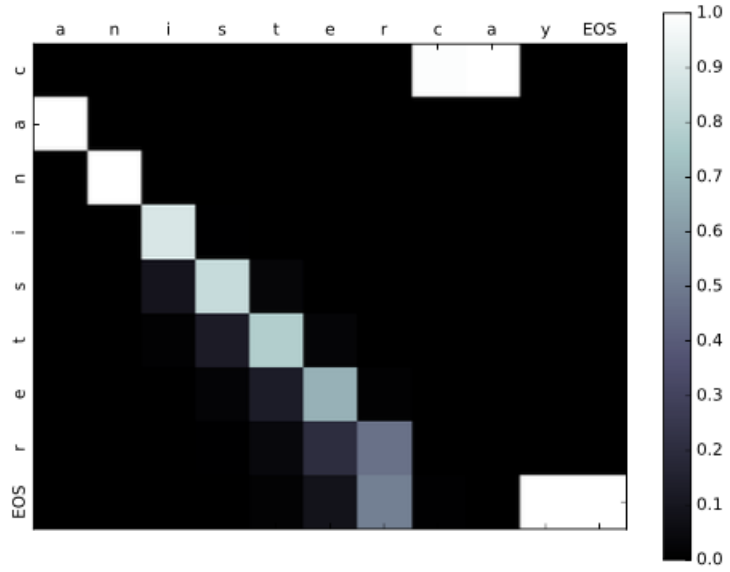

Figure 4: working: canister - anistercay

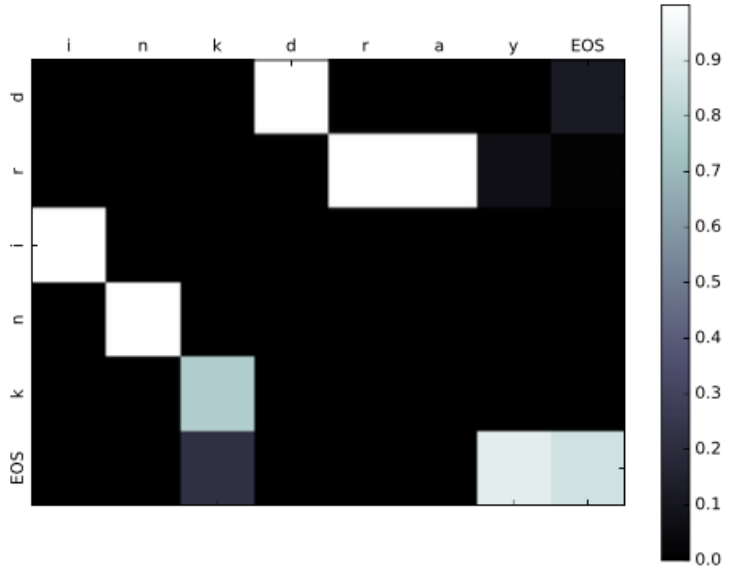

Figure 5: working: drink - inkdray

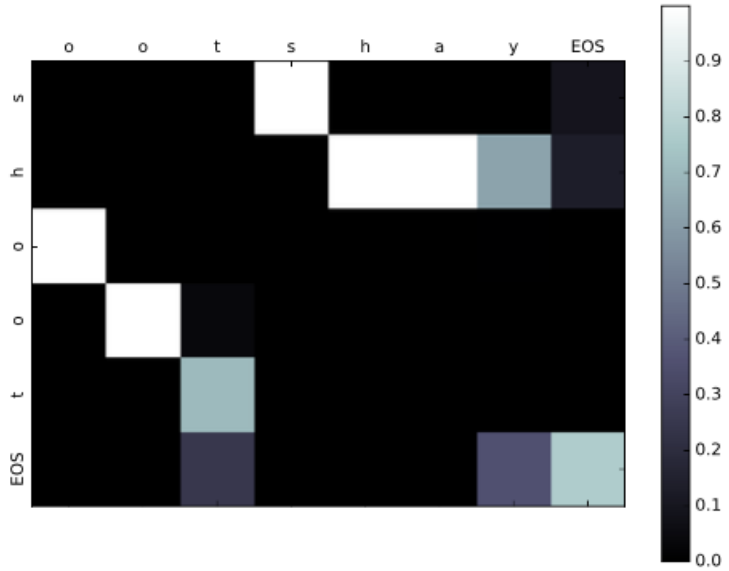

Figure 6: working: shoot - ootshay

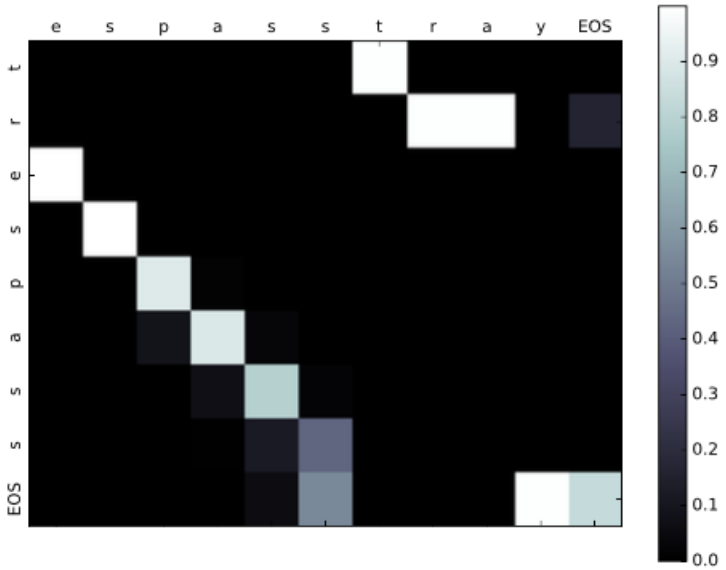

Figure 7: working: trespass - espasstray

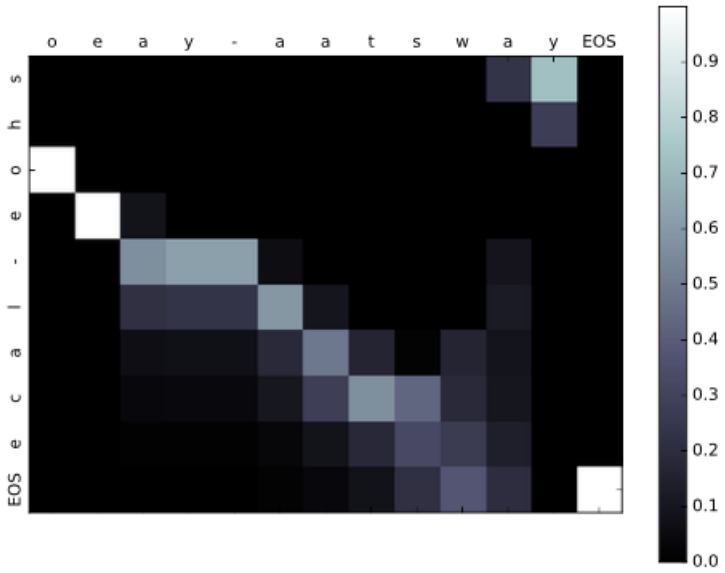

Figure 8: NOT working: shoe-lace - oeay-aatsway

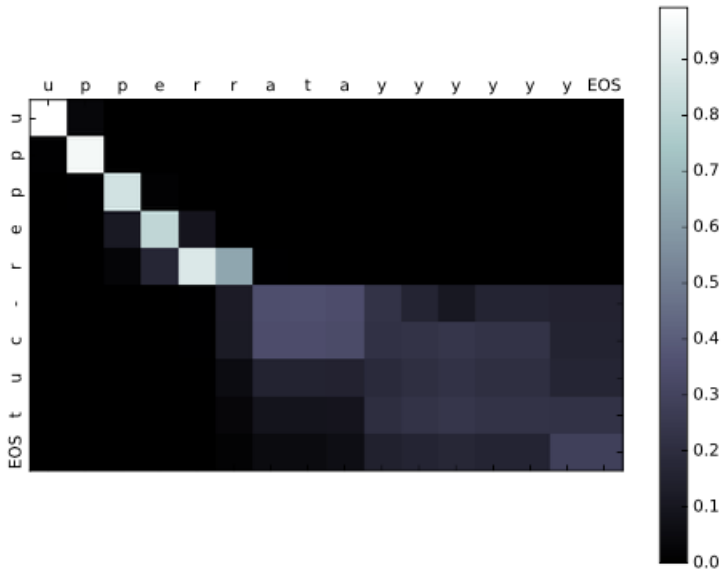

Figure 9: NOT working: upper-cut - upperratayyyyyy

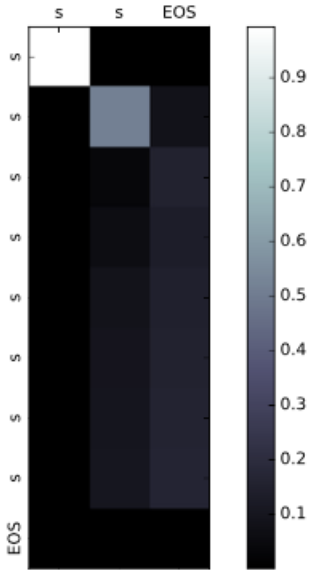

Figure 10: NOT working: ssssssss - ss

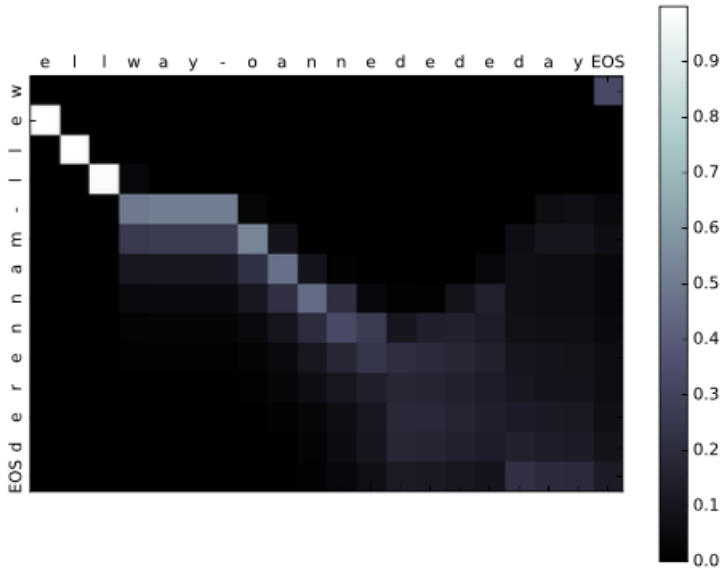

Figure 11: NOT working: well-mannered - ellway-oannede